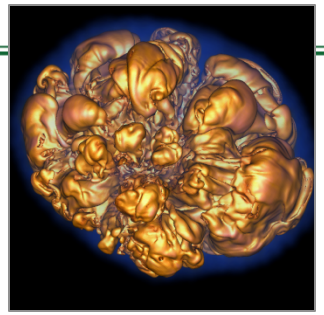
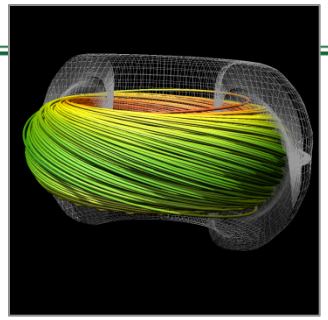
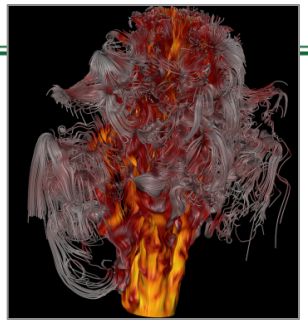




Secrets of the Elusive "Super-Mega" X/Y Chart: Snapshot of High Performance Scientific Visualization Research

E. Wes Bethel

4 October 2011

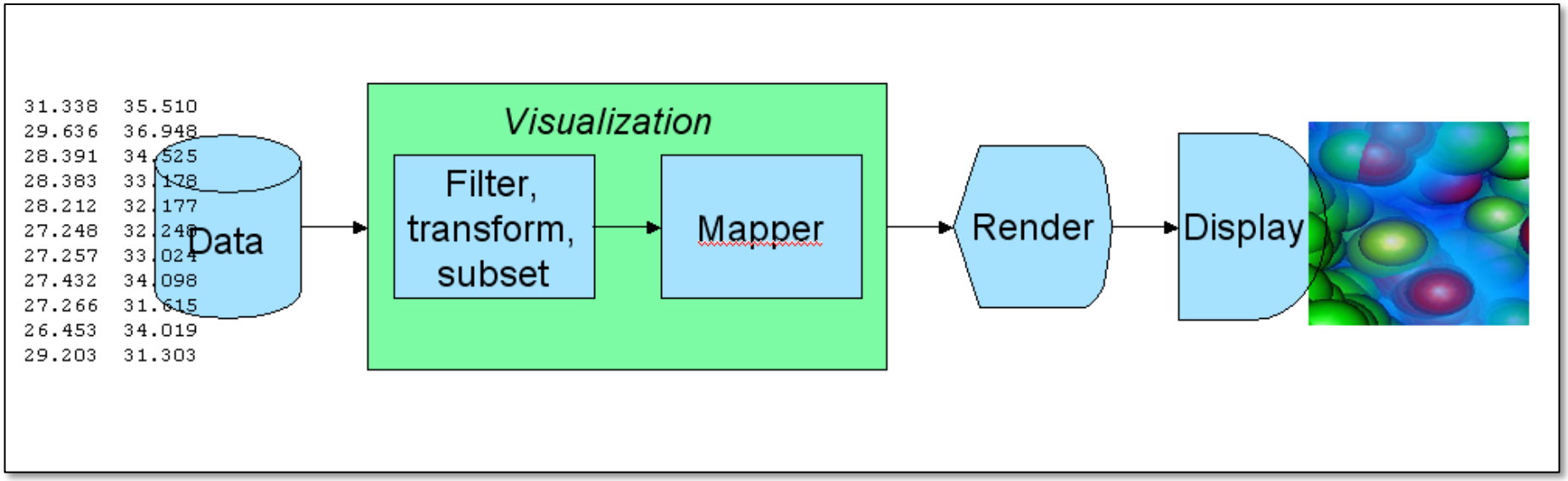
ALS Users Meeting

Berkeley, CA, USA

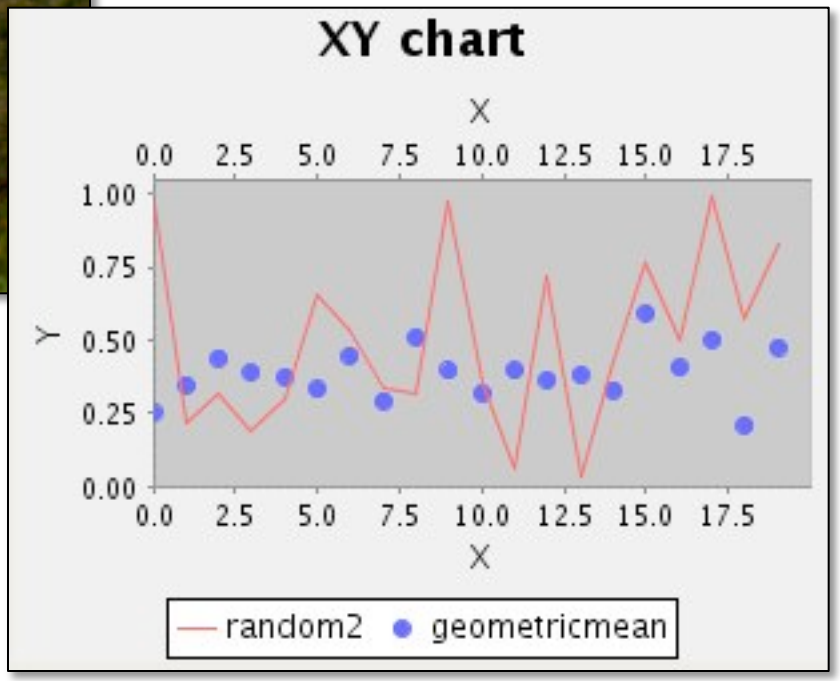
U.S. DEPARTMENT OF ENERGY | Office of Science

What is Visualization?

- Visualization: transformation of data into images.
  - Analysis: transformation of data into numbers?
- Problem: Moore's law growth of data:
  - $f(P,n,D_c,S) \rightarrow D_t$
- The purpose of visualization, analysis is insight.



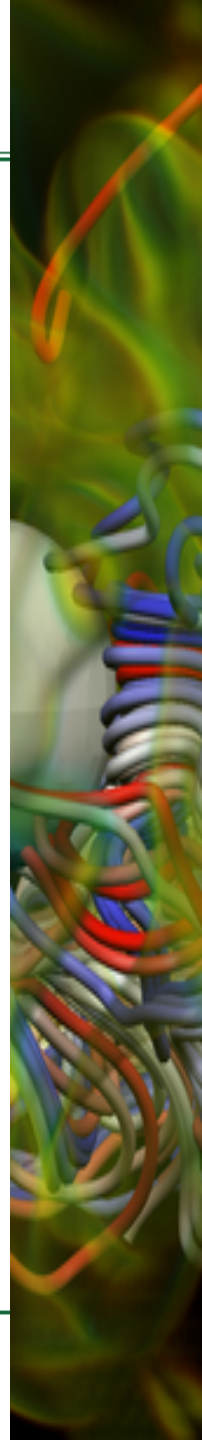
The Super-Mega X/Y Chart



$ a.out | sort - | gnuplot - // won't work

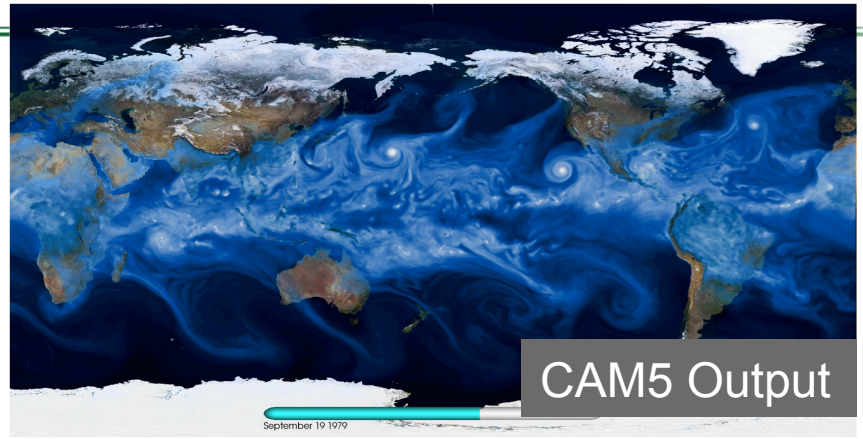
Some Current Visualization Research Issues

- Coping with increased data size and complexity in very question-specific ways.
  - Case studies from combustion, climate, and accelerator modeling.
- How to effectively utilize emerging computational platforms?
  - Extreme concurrency on multi-core platforms.
  - Using an extended memory hierarchy to accelerate vis/analysis.

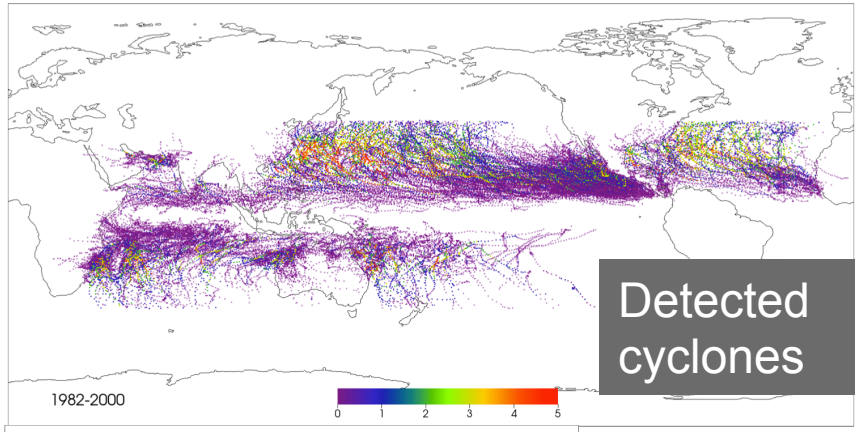


Cyclone Detection

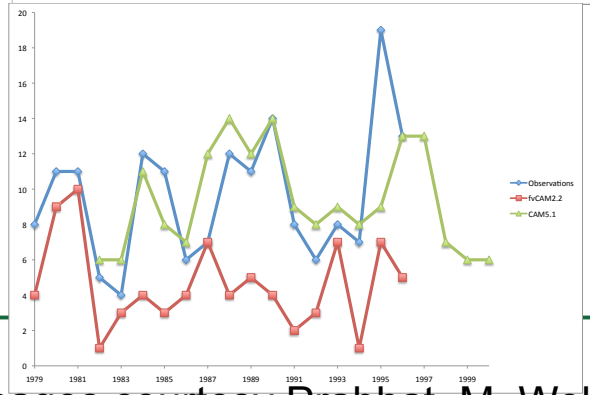
- Science objective: quantify hurricane/cyclone characteristics in a changing climate.
- Recent CAM5 0.25° runs at for 1982-2000 result in 100TB of model output.
- GFLD tracking code parallelized over time.
- Analysis time:
  - 2hrs wallclock on 7K CPUs.
  - Est. serial time: 583 days.



CAM5 Output



Detected cyclones



Annual cyclone counts

Images courtesy Prabhat, M. Wehner (LBNL)

Topological Analysis of Large-Scale Data

- Science question: what is the relationship between turbulence and combustion characteristics?
- Approach: want a quantitative basis for measuring, comparing.
- Solution: topological segmentation, analysis of combustion model output leads to “super mega X-Y chart” showing answers.

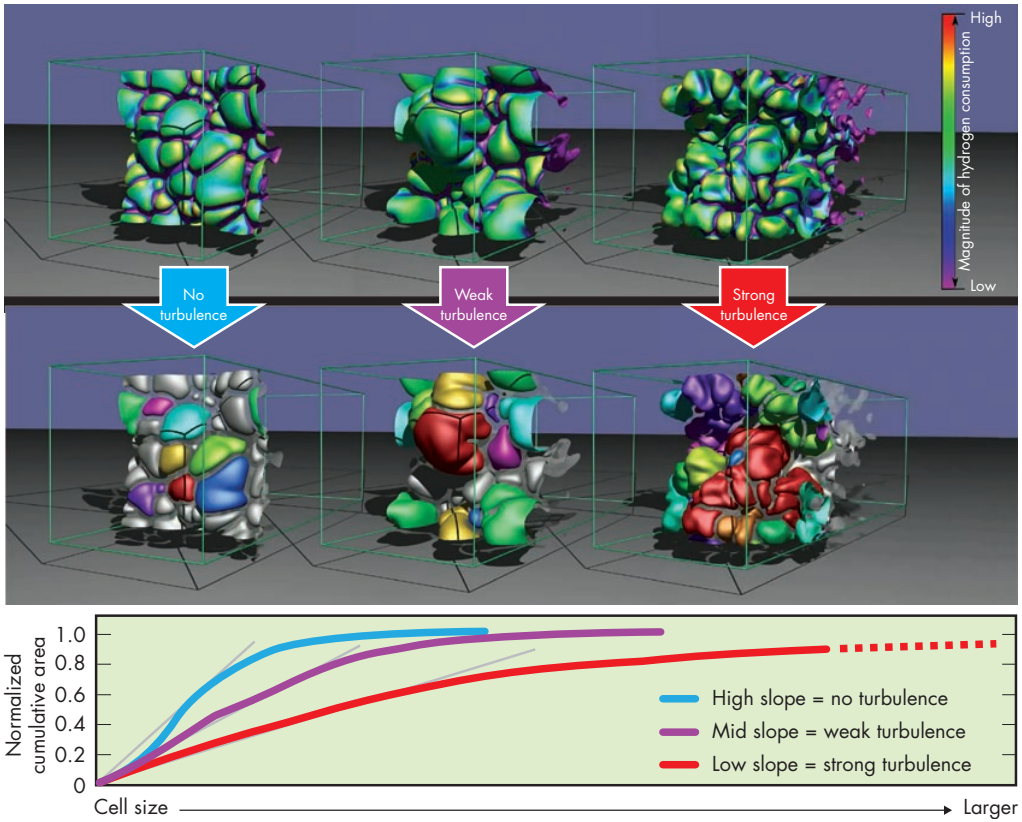
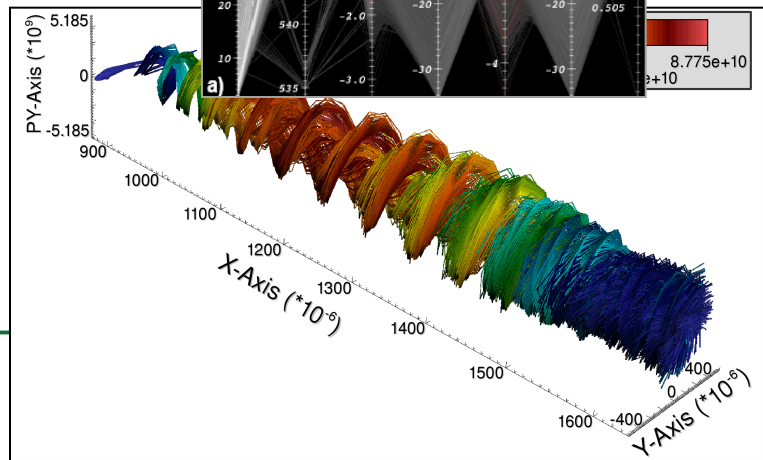


Image courtesy: V. Pascucci (Utah), P-T Bremer (LLNL), G. Weber (LBNL), J. Bell & M. Day (LBNL)

Accelerator Modeling

- Accomplishment:
  - Algorithms and production-quality software infrastructure to perform interactive visual data analysis (identify, track, analyze beam particles) in multi-TB simulation data.
- Science Impact:
  - Replace serial process that took hours with one that takes seconds.
  - New capability: rapid data exploration and analysis.
- Collaborators:
  - SciDAC SDM Center (FastBit)
  - Tech-X (Accelerator scientists)
- PI: C. Geddes (LBNL), part of SciDAC COMPASS project, INCITE awardee



Hybrid-Parallelism and Extreme Scale Visualization – E. Wes Bethel (LBNL)

ASCR- Visualization Research Highlight

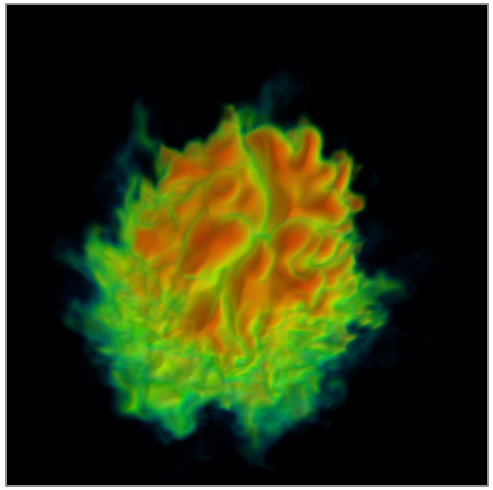
Objectives

- Fundamental R&D to enable visualization at the exascale.
- Study alternative formulations of parallelism of staple visualization algorithms.
- **Gain better understanding of limits to extreme scale visualization/analysis algorithms.**

Impact

- Results show that evolution from current approaches is required to reach the exascale for visualization, suggests fruitful paths for future research and development.
- Highest levels of concurrency ever published in the field of visualization, Best Paper Award at Eurographics 2010.

Hybrid-parallel volume rendering of 64 billion zones (300TB memory footprint) from combustion simulation output on 216,000 cores of JaguarPF.



Progress & Accomplishments (FY10 & FY11)

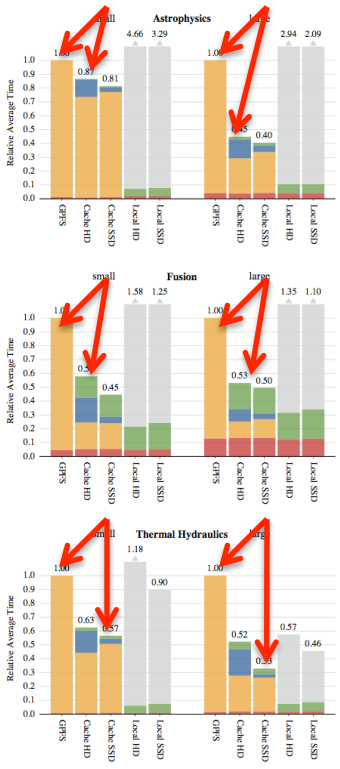
- **Results show the hybrid parallel approach: runs faster, consumes significantly less memory at all stages of processing, requires less data movement.**
- **Strong and weak scaling studies on jaguarpf, and distributed memory GPU systems.**
- **Best Paper Award at Eurographics 2010, multiple conference proceedings papers, IEEE Transactions on Visualization and Computer Graphics journal article, 1 PhD dissertation.**

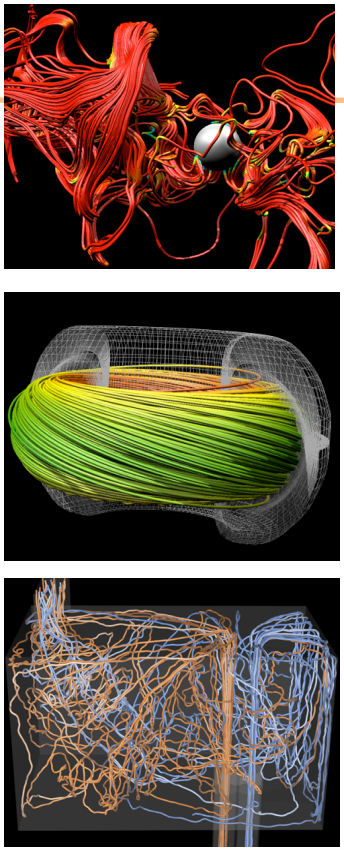
LBNL Visualization Base Program Study on Using Future Architectures: David Camp, Hank Childs, E. Wes Bethel (LBNL)

ASCR- SciDAC/Visualization Highlight

Objectives

- We must understand how to best use future ASCR computers, which may have dramatically different architectures.





Impact

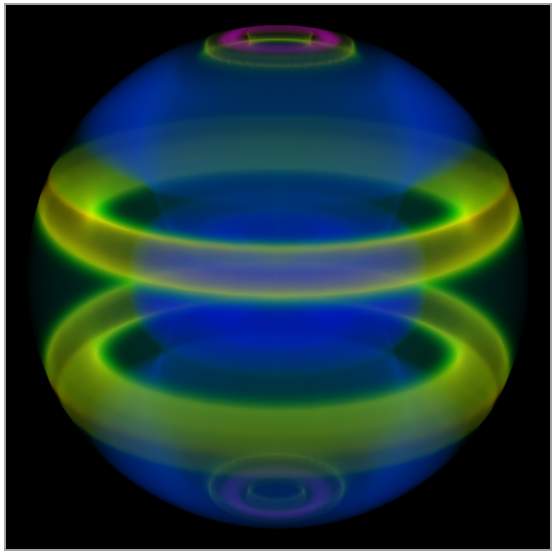
- We found that the performance of a common visualization algorithm (streamline performance) can be improved by up to *factor of three* when making use of the extended memory hierarchy likely to be commonly available on future machines.

Progress (and/or Accomplishments w/FY)

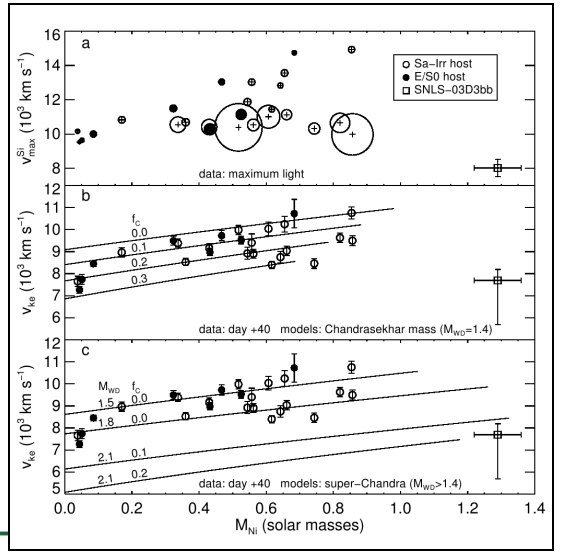
- **Designed and implemented streamline generation algorithm to make use of solid state drives (SSDs). Repeated reads of data (and their subsequent purges) are avoided by treating SSDs as a large cache. (FY11)**
  - **Although SSDs on exascale machines are designed to accelerate *writes*, we wanted to see if algorithms that perform frequent *reads* could also benefit.**
- **Performed study on SDSC Gordon machines, which has a FLASH drive on every node. (FY11)**
- **Paper accepted to IEEE Visualization 2011 Workshop on Large Data Analysis and Vis (LDAV)**

Additional Thoughts

- Secrets of the “Super-Mega X/Y Chart”
  - Clean, consise presentation of some idea.
  - Climate example: 100TB reduced to a single x/y chart.



- Three visualization use modalities:
  - Exploratory, analytical, presentation.
  - Different tools, techniques for different needs.



The End

- More information: http://vis.lbl.gov/, ewbethel@lbl.gov.